

Entrepreneurial discrete choice: Modelling decisions between self-employment, employer and worker status

Robert J. Bennett, Piero Monteburuno, Harry Smith, and Carry van Lieshout

rjb7@cam.ac.uk pfm27@cam.ac.uk hjs57@cam.ac.uk cv313@cam.ac.uk

Working Paper 15:
Working paper series from ESRC project ES/M010953:
Drivers of Entrepreneurship and Small Businesses

University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure, Downing Place, Cambridge, CB2 3EN, UK.

February 2019

[An earlier version of this paper was presented at the ISBE Annual Conference in Belfast in November 2017]

Comments are welcomed on this paper: contact the authors as above.

© Robert J. Bennett, Piero Monteburuno, Harry Smith, and Carry van Lieshout, University of Cambridge, members of the Cambridge Group for the History of Population and Social Structure assert their legal and moral rights to be identified as the authors of this paper; it may be referenced provided full acknowledgement is made: *Cite* (Harvard format):

Bennett, Robert, J., Monteburuno, Piero, Smith, Harry, and van Lieshout, Carry (2019) *Entrepreneurial discrete choice: Modelling decisions between self-employment, employer and worker status*. Working Paper 15: ESRC project ES/M010953: 'Drivers of Entrepreneurship and Small Businesses', University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure.

Keywords: Entrepreneurship, Employers, Self-employment, Small businesses, Census

JEL Codes: L26, L25, D13, D22

Entrepreneurial discrete choice: Modelling decisions between self-employment, employer and worker status

Robert J. Bennett, Piero Monteburuno, Harry Smith, and Carry van Lieshout

Working Paper 15: ESRC project ES/M010953: Drivers of Entrepreneurship and Small Businesses, University of Cambridge.

1. Introduction

This paper estimates an entrepreneurial choice model between different forms of business organisation, focusing on the choice to operate as a sole proprietor operating on own account, or as an employer, rather than as a waged employee. ‘Own account’ was the phrasing used through the censuses of the period we cover: it describes an entrepreneur who operates as self-employed on their own with no employees. A discrete choice decision model is used to estimate probabilities of different choices using the large scale data for England and Wales contained in the ‘British Business Census of Entrepreneurs 1851-1911’ to be deposited at UKDA. This paper uses the data for the later censuses, 1891-1911. This database is one of the outputs of ESRC project ES/M010953 Drivers of Entrepreneurship and Small Businesses’. The data referred to in this working paper for 1851-1881 is derived from the Integrated Census Microdata (I-CeM) deposited at UK Data Archive (UKDA), which has been used in a revised and updated form. The I-CeM records are derived from the transcriptions made by the commercial genealogy provider Find My Past (FMP) (part of BrightSolid) in conjunction with The National Archive (TNA).

The process of identification and extraction of the data on entrepreneurs from the original census household returns is described in Working Paper 4, with the methods of weighting and

adjustment of the census returns for 1891-1911 to correct for non-response and misallocation bias described in Bennett et al. (2019), and weights given in Montebruno (2018). Working Paper 2 defines outlines the different census questions and the challenges they present for identifying entrepreneurs. An overview of the research strategy and full data assembly process underpinning the database for entrepreneurs is given in Working Paper 1. A full list of Working Papers is included at the end of this paper.

An important part of the interface between historical and modern debates on entrepreneurship is recognition of the important role of individual characteristics, demography and family structures. These appear to be enduring influences on entrepreneurship across time. In modern research Howarth et al. (2010), Carter and Ram (2003), Colli (2003), Colli et al. (2013), Alsos et al. (2014) and many others have focused on the changing role of households and families in small business development. In historical research Anderson (1971, 1988), Church (1993), Nenadic (1993), Davidoff and Hall (1997), and Davidoff (2012), among many others, indicate that family, kinship, inter-generational ties and co-residence were key features of business development in the nineteenth century. Distinctive features of demographic and family structures that benefit family firms and single-person own account businesses usually rely on family support of labour, capital and other inputs, as well as wider personal networks.

However, historical analyses have often been held back by the lack of large scale data. This has limited the ability to generalise and scale up from small scale and case study research. This paper overcomes the data scale issue by using the newly available large scale electronic database of the censuses for 1891-1911. These historical censuses have similar structure to modern censuses so that examination of the 1891-1911 period lays the foundations for comparing modern and earlier decision choices and the factors that underlie them.

The remainder of the paper is organized as follows. Section 2 discusses theories of entrepreneurial choice that underpin distinctions between different statuses. Section 3 discusses the methodology, data used, and empirical strategy. Section 4 presents and interprets the empirical results. The final section concludes and assesses the significance of the findings. The estimation focuses on 13 sectors that are market-facing in which a real choice between status (of employer, own account proprietor, or worker) operate; sectors that

were solely workers are excluded (public administration, military), as well as employment of domestic staff.

2. Theories of entrepreneurial choice

Our focus in this paper is entrepreneurial decision-making between different forms of business organisation: to operate as a sole proprietor or to employ others, as distinct from being a waged employee. Self-employment without employees can be an attractive choice allowing independence, but is often associated with necessity and survival entrepreneurship. In the history of the period we examine this is referred to as operating on ‘own account’. Taking employees into a business often marks a step to a larger and more sustained business, but can be more cumbersome and off-putting because of higher transaction costs for internal management and decision making than operating alone. Of course, such choices also interact with opportunity, which may be constrained, so that both supply and demand issues must be investigated.

Much of the literature on business organisation investigates choice between proprietorship and the corporate form. We are not concerned here with the issues of incorporation, which is an important dimension which we reserve to further analysis; it is somewhat separate from our focus on choices between sole proprietorship and employing others. However, the literature on the corporate form provides some important indicators of the theoretical choice framework required. For Chandler (1962) incorporation underpinned hierarchical management that allowed internal economies of scale and scope, whilst for Williamson (1975) these could be broken down into various transaction cost advantages of the corporate form. Jensen and Meckling (1976) have developed these concepts to focus specifically on entrepreneur/manager choices and principal-agent relations of those inside and outside the firm. This has stimulated an important literature comparing different legal forms of incorporated businesses focusing on shareholder and corporate financial controls (La Porta et al., 1998, 1999, 2008). However, Jensen and Meckling (1976, p. 311) rightly note that what matters most are the actual relationships between the individuals concerned, and these apply as much to the choices between sole proprietorship and employing others as to choice of incorporation: the monitoring costs of principal (owner-manager or entrepreneur), bonding costs of agents (co-entrepreneurs and employees), and residual transaction costs (and losses).

However, for many small enterprises, and all those considered here, the issues of choice is whether to extend sole proprietorship working on one's own to a business employing others. As noted by Pollard (1965) and Ang et al. (2000), sole proprietors with no employees are a base case of zero agency-costs, as specified by Jensen and Meckling (1976). The crucial step of taking on employees indicates decisions to engage in business expansion, but also to take on higher management and other transaction costs, as well as bonding costs and the challenges of managing agency effects.

One of the first theoretical contributions to focus specifically on choice between sole proprietorship and employing others was Knight (1921). He argued that people respond through choices that balance uncertainty against their personal characteristics, ability, opportunity, information available to them, and other factors. Those individuals favouring higher certainty generally prefer waged employment that relies on others to face the challenges of running a business. Those favouring greater independence and willingness to take risks will be more willing to become entrepreneurs. Knight's central theoretical argument was that uncertainty shapes choices through: the extent to which people have adapted to particular occupational situations, and developed particular skills and experiences of making judgements; as well as their ability, experience, and ability in risk taking (Knight, 1921, p.270).

We also recognise that the constraints of agency work both ways for choices. For employees, salaried status was/is often preferred and in the Victorian period the preference for waged employment was often stimulated by the relatively high pay available to men, combined with increasing social convention, which led to a dominance of the male bread-winner family. Supply also influenced the availability of suitable business partners or other supports: because of age, life cycle, or available family and networks individuals varied in their access to others with whom they could share supports to develop as entrepreneurs. As well as individual contexts, the wider locational and economic context will also be important: particularly the labour market, the overall population numbers and its composition, especially locally; as well as access to other agents controlling factor inputs (banks and finance institutions, advisory expertise, etc.).

Sectors are an important aspect of choices and are a core element in our analysis. Some sectors have much lower entry barriers than others favouring start-ups and individual self-

employment. This will be especially important to necessity entrepreneurs. Also, some sectors are only scalable by adding more personnel, which is the case with most service industries. The choice model sees independent own account proprietors as favoured in sectors that have low entry barriers and are difficult to scale up. Or if scalable, can be managed through easy aggregation of small numbers of personnel, especially if possible through divided roles between individual partners and/or family and close networks. This can range across all sectors, except those where large scale production is essential e.g. large scale steel ship building or iron and steel production. But sole proprietorship is most favoured where higher skills are required, and/or in specialised fields with high knowledge levels. Fields especially favoured are those where the knowledge is not easily aggregable and sub-dividable. This characterised specialist manufacturers in craft industries (such as watch and instrument making), as well as many professions, such as specialist engineers, architects, doctors and lawyers. Similarly, if the only way to scale up is by increasing personnel, small firms and individual self-employed can often compete effectively on quality and/or price; as in care industries, much retailing, lodgings, and other services. In the nineteenth century it also characterised industries where large scale factory manufacture was less able to compete with the individual, such as many small scale artisan manufactures (jewellery, decorative arts and craft industries, instruments, watch and clock making), sectors where small manufactures could compete on quality or some types of product specialism (shoes, clothing, many food manufactures), many building and construction trades (painters, plasterers, carpenters, bricklayers), washing and laundry (even after large scale steam laundries began to take over), and local retail, merchanting and trading. Many of these sectors have been regarded as ‘traditional’ industries compared to those where factories and corporations had most strongly developed, or in modern time where electronic trading is possible, but many of these remain an important parts of historic and modern small business and sole trading activity.

3. Methodology and estimation

This paper is a national level analysis of entrepreneurs identified as either self-employed sole proprietors, employers of others, or waged workers for England and Wales. These are identified and extracted from the 28-40 million records within the population censuses 1891-1911. The data provide the opportunity to develop for three historical years a method of choice modelling that can be extended and compared in the future to other census years and

data. The database used has become available through encoding the information in the original manuscript census records. The data referred to here for 1891-1911 are derived from the Integrated Census Microdata (I-CeM) deposited at UK Data Archive (UKDA) (Higgs et al., 2015; Schürer et al., 2016), which after cleaning and coding for entrepreneurs are available in the ‘British business Census of Entrepreneurs 1851-1911’. The data on individual entrepreneurs used here derive from the original Census Enumerators Books (CEBs) for 1891 and 1901, and from the original householders’ returns for 1911.

The focus of this paper is estimation of an econometric discrete choice model for the employer and sole proprietor decision compared with waged employment at the level of each individual, controlling for interaction with other factors. Not all desirable control variables are available for this historic period; for example, education level used in many modern entrepreneurship studies is not available at individual level. However, the main range of the controls usually adopted in entrepreneurship studies where the focus is on personal characteristics can be included: co-resident family and others; marital status, gender; age; business sector; and geographical location. Appendix Table A1 lists the definitions and characteristics of each variable included. The sector definitions are derived by aggregating the 797 occupational classifications used in the censuses into 13 groups that contain entrepreneurs. The definitions of the groups are given in detail in WP 5. They are a development from contemporary and modern research census research (see Booth, 1886; Armstrong, 1972). In contrast to some other studies, there is no attempt to separate all manufacturers from dealers (and hence secondary and service sectors) where this is not possible because of the way the census data was collected, but instead a category of ‘maker-dealer’ is used for the many small businesses, such as shoe makers, dressmakers, tailors or bakers, that manufactured but also retailed directly. This was a major characteristic of much manufacturing in this period.

Estimation of decision choices is undertaken through two models. In **Model 1**, which is the core decision model between business organisational choices, the dependent variable is a multinomial indicator in three categories: employer, own account self-employed, and worker. Explanatory variables include age, population density and number of servants as continuous variables. Age is measured as both a linear and squared term to capture expected nonlinearities. Population density is measured per acre for each Registration Sub-district (RSD) (there were 3,000 of these in 1891-1911). Population density directly measures the

level of urbanisation, but also controls for potential non-linear relationships with entrepreneurship indicating changing opportunities at different levels of localised market potential. The model also includes a variable to capture the effect of number of servants. Although we are unable to control for entrepreneurs' education and abilities, we have information on a form of social capital, the number of servants which indicates something of the wealth and other resources available that may release entrepreneurs' time as well as in some cases contributing to the business, where it is known that in many Victorian family businesses domestic servants also assisted in the business; e.g. in shops and on farms. Initially, in Model 1, marital status and gender are combined to create six dummy variables representing single men, single women, married men (the base category), married women, widows and widowers. The data contain 13 sectors, estimated as dummy variables (see Table A1; the base category is farming). It should be noted that the 13 sectors are those where an entrepreneurial choice operates. There are 4 other sectors of activity which are solely waged and hence contain exclusively workers (public administration, military, clergy; domestic and other home service staff; undefined general labourers; and persons of property with no stated occupation). This exclusion allows focus on the sectors where a real choice of employment status exists.

Model 2 seeks to test alternative relations between family and intra-household structures. It includes additional variables for household relationships by modelling 10 dummy variables for various family and non-family connections identifiable from the census. Because of the overlap of these categories, gender and marital status is now estimated separately between the different household relationships. The census question required each person staying in the household on the night of the census to be recorded with their relationship to the 'head' of the house. The head was the census respondent and expected to be the most senior householder. This was normally the husband in a married household with children, but in more complex or simpler households could be a wife, single man or woman. Where the premises was an institution, the head would be the live-in proprietor (or resident manager) who would return themselves as well as those staying there (such as those lodging in hotels, resident staff and pupils in schools, or resident staff and patients in hospitals). The way the range of relatives and others that were then described can only be understood and coded with reference to the head; for example the wife or husband, children, grandchildren, grandparents, uncles, step and in-laws, as well as others living in the house at the time: domestic live-in servants, staff and assistants of a business proprietor living in the same house, boarders and lodgers, other

relatives and visitors, and the staff and inmates of institutions (see Table A1; the base category is head, and the base categories for gender and marital status are male and married). Other explanatory variables correspond to Model 1. Both models are estimated below.

It should be noted that in both models, employers of domestic staff, which were common in Victorian times, are excluded by the census as being an employer. This is in complete contrast to the modern census and modern British small firm statistics that count all employers equally. We have not attempted to identify these employers and include them in our analysis since the focus is on market-facing economic activity; but future analysis can estimate the decisions choices affecting the domestic service sector using the database deposited.

These models develop directly from an earlier analysis that used multinomial logits to estimate the choice model behind portfolio businesses (Radicic et al., 2017). A similar modelling approach is developed below to facilitate comparisons. Our unit of analysis is individuals. But for some variables, spatial data are used at the level of Registration Sub-Districts (RSDs). This means that people in the same geographical cluster are correlated, as they share common cluster-level random effects or level-2 error, u_j , accrued from common population density, markets and culture. This is taken into account in the estimation. Equation (1) represents level 1 model (in our case at the level of individuals) with a random intercept (Rabe-Hesketh and Skrondal, 2012):

$$y_{ij} = \beta_{0j} + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_n X_{nij} + e_{ij} \quad (1)$$

where y_{ij} is a binary response variable for individual i in a sub-district j taking a value of 1 if the individual is an entrepreneur and 0 if not, a set of coefficients to be estimated are denoted by $\beta_{0j}, \beta_1, \dots, \beta_n$; $X_{1ij}, X_{2ij}, \dots, X_{nij}$ is a set of n covariates or explanatory variables, and e_{ij} is the random effect or level-1 error. The random intercept characterizes Equation (2) below and represents level 2:

$$\beta_{0j} = \beta_0 + u_j \quad (2)$$

We need two assumptions to identify the model. First, a mean-independent assumption is required: $E(e_{ij} | \mathbf{X}'_{ij}, u_j) = 0$, where $\mathbf{X}'_{ij} = (X_{1ij}, \dots, X_{nij})$ is a vector of the n covariates

described previously. This assumption implies $E(e_{ij}|\mathbf{X}'_{ij}) = 0$ and $Cov(\mathbf{X}'_{ij}, e_{ij}) = 0$, this last corresponding to individual or level-1 exogeneity. Second, another mean-independent assumption $E(u_j|\mathbf{X}'_{ij}) = 0$ implies $Cov(\mathbf{X}'_{ij}, u_j) = 0$, which corresponds to registration sub-district or level-2 exogeneity.

The model can also be expressed through a latent variable conceptualization (Guo and Zhao, 2000). If y_{ij}^* denotes a latent variable such that $y_{ij}^* > 0$ when $y_{ij} = 1$ and $y_{ij}^* \leq 0$ when $y_{ij} = 0$, then the model for a latent variable y_{ij}^* can be written as:

$$y_{ij}^* = \beta_0 + \beta_1 X_{1ij} + \beta_2 X_{2ij} + \dots + \beta_n X_{nij} + u_j + e_{ij} \quad (3)$$

Where u_j is the Sub-District or level-2 error term and the e_{ij} is the individual or level-1 error term.

Given the nature of the data defining the dependent variable in both Models 1 and 2, which have to be corrected for non-response and misallocation biases (Bennett et al., 2019), we use a weighted estimates (Monteburno, 2018): a weighted logit model to assess the probability of being any type of entrepreneur compared to a worker; and a weighted multinomial logit to assess the separate probability of being an employer, or own-account entrepreneur, compared to being a worker. We use multinomial logit and not conditional logit because all our covariates are state-invariant regressors; that is we do not have different *ages*, *densities* or *marital statuses* for different employment statuses.

The general equation for a multinomial logit (Cameron and Trivedi, 2005) is as follows:

$$p_{ij} = Pr[y_i = j] = \frac{\exp^{x_i \beta_j}}{\sum_{j=1}^m \exp^{x_i \beta_j}}, \quad j = 1, \dots, m \quad (4)$$

Where y_i is the employment status of each individual i , and j has three values (*worker* = 1, *employer* = 2 and *own-account self-employed* = 3). Thus $m=3$ in our particular model. For this particular case, but also for the general case with any m , $Pr[y_i = 1] + Pr[y_i = 2] + Pr[y_i = 3] = 1$; ($Pr[y_i = 1] + \dots + Pr[y_i = m] = 1$); i.e. all probabilities sum to one.

In the multinomial logit model, the estimated coefficients do not imply an absolute or direct effect on the probability of being in a particular j state. They should be interpreted as a relative increase in the probability of being in a particular state with regards to the $j=1$ state or base outcome. For example, a positive coefficient for a variable does not mean that it

increases the absolute or direct probability, but only that it increases the likelihood, e.g. of being an employer relative to being a worker, the base outcome. Some authors prefer to use a relative risk or log odds ratio to assess this effect or relative magnitude but this is not used here.

Weighted estimation responds to the constraints in the data. Despite the potential of the census data used here, limitations have to be recognised and managed. The census as a source has the advantage of national coverage, collected under a legal obligation for people to reply honestly, and had significant administrative effort to ensure consistency. However, it was not designed as a business census, which results in important constraints that have to be managed in the estimation.

The census identifies self-employed ‘own account’ and employers directly through a specific question, used for the first time in 1891, which has been continued in a modified form into the modern census. In 1891 the question instructed respondents to put a cross in one of three columns headed ‘Employer’, ‘Employed’, and ‘Neither Employer nor Employed’. The last category was expanded in the instructions, but not over the column, as ‘the person neither employs other workmen in his trade or industry, nor works for a master, but works on his *own account*’. This survey design was defective in various respects, particularly in the similarity of ‘employer’ and ‘employed’, and the negative header definition of own account. The gendered language, though typical of the time, was also potentially distortionary, which was exacerbated by a further instruction about wives who were partners or co-preneurs: ‘Married women assisting their husbands in their trade or industry are to be returned as “Employed”’.¹ Because of recognised difficulties in interpreting the results of this question at the time, it was modified in 1901 to ask respondents to write in their occupational status as “Employer” (that is, employing persons other than domestic servants), “Worker” (that is, a worker for an Employer), or “Own account” (that is, neither Employer nor working for Employer, but working on own account)’.² Almost identical wording was used in 1911 which has remained similar up to the present. The 1901 and 1911 instructions generally produced sound and reliable responses. But the deficiencies of the 1891 instruction mean that various adjustments have to be made for mis-allocation biases between own account and employers, and between workers and both categories of entrepreneur.

¹ ‘General Instruction’, Census of England and Wales, *Householder’s Schedule*, 1891.

² ‘General Instruction’, Census of England and Wales, *Householder’s Schedule*, 1901.

In addition, for all three years there was a relatively high level of non-response to this particular question. There was (and is) a legal obligation to reply to the census, but the 1891-1911 administrators of the process did not always check that all questions were responded to, and this question was deemed of lower priority by census administrators than most other questions (on occupation, age, gender, etc.). Moreover, there were significant biases to the non-responses: higher in some sectors than others, higher for women, and also higher for individuals within households other than the ‘head’ (the person who filled in the census form), (WP 4; see also Bennett et al., 2019). Much of the non-response bias is corrected by data cleaning to remove those who ticked an entrepreneur status who were scholars, under 15, were economically inactive, or in worker only categories (domestic service, labourer etc.). The remainder were re-weighted taking account of occupation, gender, age, marital status and position in the household. The weights are available as a data download (Monteburno, 2018).

Although data weighting is an important exercise it is important to bear in mind that comparison of the estimates for weighted and unweighted results show no differences in significance levels or coefficient sign values for any of the variables, and the change in coefficient values was inconsequential.

4. Empirical results

4.1 Model 1: Estimates for joint categories of marital status and gender

The results of estimation of the model using Equation 1 are reported in Tables 1-3 for 1891, 1901 and 1911, respectively. In these tables worker status is the base category and the model estimates the probability of being either an employer or own account self-employed. In each table two different models are estimated and compared: in column 2 the weighted logit model of the probability of being *either* an employer or an own account self-employed is estimated; i.e. all entrepreneurs compared to being a worker. In columns 3 and 4 the weighted multinomial logit model of the probability of *separately* being an employer or an own account self-employed is estimated compared to being a worker. This allows comparison of all entrepreneurs against all non-entrepreneurs at the date of the census. The estimation in column 2, by combining employers and own account, also allows any remaining errors in

classification between entrepreneur categories to be overcome.

Focusing initially on the impact of gender and marital status on the probability of being an employer (column 3 of Tables 1-3), the estimates suggest that all categories of individuals were less likely to be employers than married men (the base category): with single women the least likely of all to be employers, followed in declining rank order by married women, and then single men, widows and widowers. This strongly confirms the dominant role of the Victorian marital unit as a major advantage in mutual resource support: freeing the husband to develop the business whilst wives and family provided household support and other unpriced inputs to the family business. However, when death of the spouse occurred, women as widows showed higher probability of maintaining or initiating employer status than men as widowers. Such widows may also have already been de-facto partners in the business but this was not recorded in the census response.

For own account entrepreneurs (column 4 of Tables 1-3) the pattern is very different. Married men no longer had the highest probability of entrepreneurial activity; instead it was married women, followed by widows, and single women, each of which were more probable to be own account than married men. Single men were less likely to be own account than widowers, who were both less likely to be own account than married men. This is a striking and significant comparison not previously noted in 19th century studies at the level of the whole population. It confirms the general and dominant status of the so-called 'male breadwinner family' as waged labour, and also the dominance of married males among employers. But it also confirms that women were important complimentary supports to the male waged breadwinner developing income opportunities through their own entrepreneurial activities. It has often been observed that women 'earned a bit on the side'; see for example the case studies of Davidoff and Hall (1991) and Davidoff (2012). But rather than being a minor or part time occupation that wives and other women engaged in, the census (by identifying only full time activity) shows that it was a major contribution to family incomes for married women. It also confirms case studies that show widows and single women using own account self-employment as a major way they could maintain their independence. In contrast widowers and single men were much more likely to go into waged worker status.

Taking entrepreneurs as a whole (column 2 of Tables 1-3), there is an averaging, with the higher numbers of own account than employers weighting the estimates to the greatest extent.

But the surprising result is that the highest probability of being an entrepreneur was for widows, then married women, which were all greater than for married men. Single women had almost the same probability of being an entrepreneur as married men, followed by widowers, and lowest of all for single men. The modern literature tends to have focused on the growing levels of entrepreneurship among modern women, but it is clear that this was a phenomenon of long standing, and that modern trends may be only recovering levels of female self-employment that were much more widespread in the 1891-1911 period than in many of the intervening years. The late Victorian and Edwardian period appears in many ways as an age of female entrepreneurship. Of course many of their activities were support to family incomes, and hence may have done no more than operate as survival and necessity entrepreneurs, as indicated by the dominance of own account female activities in dressmaking, laundry, inn keeping and shops. These were usually limited in scale of activity with less scope to grow, and utilised predominantly traditional craft and hand skills. However, many were in sectors offering significant scope for scale, especially in some manufacturing; and were also becoming prominent in professions, especially as school proprietors. Hence, female entrepreneurship was more significant than often contended and existed more strongly than for men in some sectors and locations, especially where the dominant pull of the higher waged opportunities and the greater security of wage employment frequently tilted towards recruiting males. It is also to be borne in mind that with the dominance of males in the waged labour market the main alternative for women was domestic service, so that female entrepreneurs were those who wanted to avoid this poorly paid and more menial occupation. Note that the estimation covers only those sectors that had employers, own account and workers; those sectors that were solely workers were excluded. This means that one of the well paid categories of male employment in public administration is not included, and the most frequent category of female employment in domestic service is excluded. The model focuses on the sectors where there was a choice. In those sectors a woman with independent spirit and opportunity could develop own account trades as a viable and often more lucrative alternative, with greater potential for personal satisfaction and fulfilment, even taking account of in-kind benefits offered in domestic service. The results are all the remarkable since, as noted earlier, there was a systematic tendency in the census to under-record female occupations in general, and especially those of wives and other females in the most common households that had a married male as head.

Table 1. 1891: Weighted logit model for the probability of being an entrepreneur (employer + own-account) (col.2), and weighted multinomial logit model for the probability of being a worker (base, omitted), or either employer or an own-account self-employed (cols. 3 and 4). Density at RSD level; only simple interactions and no levels for Sex and Marital status (Base categories Male and Married (Sex and Marital status), and Farming/Estate Work (13 sectors).

	All entrepreneurs	Employer	Own account
Density	-0.00697*** (-135.25)	-0.00702*** (-86.91)	-0.00697*** (-120.59)
Density # Density	0.0000194*** (85.76)	0.0000214*** (62.79)	0.0000186*** (73.25)
Age	0.149*** (356.24)	0.160*** (223.57)	0.145*** (302.70)
Age # Age	-0.00118*** (-260.30)	-0.00128*** (-170.90)	-0.00114*** (-218.52)
Single men	-0.686*** (-174.12)	-1.084*** (-162.64)	-0.469*** (-101.28)
Widowers	-0.247*** (-48.65)	0 (.)	0 (.)
Single women	-0.0929*** (-23.73)	-0.571*** (-78.90)	-0.0241*** (-3.96)
Married women	-0.0930*** (-20.53)	-1.355*** (-138.58)	0.254*** (59.02)
Widows	0.218*** (47.51)	-1.081*** (-108.71)	0.234*** (47.74)
Number of Servants	0.850*** (315.12)	-0.399*** (-50.84)	0.465*** (91.21)
mining and quarrying	-2.440*** (-210.33)	1.244*** (353.64)	0.439*** (146.88)
construction	-0.0134** (-3.05)	0.105*** (17.32)	-0.101*** (-17.18)
manufacturing	-0.865*** (-213.02)	-2.301*** (-152.73)	-2.628*** (-142.43)
maker-dealer	1.160*** (317.32)	0.0175** (3.01)	-0.0341*** (-5.65)
retail	1.412*** (248.52)	-0.738*** (-134.46)	-0.915*** (-161.97)
transport	-1.109*** (-176.89)	0.566*** (104.24)	1.532*** (340.17)
prof and bus services	-0.561*** (-81.40)	0.921*** (110.22)	1.775*** (270.69)
personal services	0.432*** (85.27)	-1.626*** (-155.21)	-0.738*** (-96.69)
agric produce processing and dealing	0.403*** (44.76)	-1.091*** (-104.25)	-0.159*** (-18.57)
food sales	1.885*** (445.44)	-0.449*** (-45.69)	0.855*** (144.29)
refreshment	1.269*** (226.76)	0.469*** (41.55)	0.314*** (24.43)
finance and commerce	0.0109 (1.22)	1.153*** (184.35)	2.319*** (462.16)
Constant	-5.464*** (-552.65)	0.162*** (15.93)	1.823*** (292.65)
Observations	8,929,277	8,929,277	
Pseudo R ²	0.268	0.254	

Table 2. 1901: Weighted logit model for the probability of being an entrepreneur (employer + own-account) (col.2), and weighted multinomial logit model for the probability of being a worker, or either employer or an own-account self-employed (cols. 3 and 4). Other definitions as in Table 1.

	All entrepreneurs	Employer	Own account
Density	-0.00946 ^{***} (-174.39)	-0.00825 ^{***} (-91.58)	-0.00993 ^{***} (-163.84)
Density # Density	0.0000279 ^{***} (104.03)	0.0000249 ^{***} (55.47)	0.0000291 ^{***} (98.01)
Age	0.157 ^{***} (396.62)	0.189 ^{***} (258.83)	0.146 ^{***} (331.52)
Age # Age	-0.00124 ^{***} (-288.87)	-0.00160 ^{***} (-207.09)	-0.00111 ^{***} (-231.85)
Single men	-0.610 ^{***} (-170.77)	-0.887 ^{***} (-145.77)	-0.463 ^{***} (-111.93)
Widowers	-0.326 ^{***} (-66.13)	-0.467 ^{***} (-67.07)	-0.228 ^{***} (-38.84)
Single women	-0.0268 ^{***} (-7.54)	-1.433 ^{***} (-145.89)	0.277 ^{***} (72.46)
Married women	0.167 ^{***} (38.46)	-1.030 ^{***} (-95.86)	0.472 ^{***} (102.23)
Widows	0.222 ^{***} (49.31)	-0.284 ^{***} (-35.50)	0.400 ^{***} (82.25)
Number of Servants	0.941 ^{***} (317.86)	1.369 ^{***} (356.39)	0.553 ^{***} (172.57)
mining and quarrying	-2.689 ^{***} (-255.18)	-2.582 ^{***} (-184.71)	-2.841 ^{***} (-176.48)
construction	-0.398 ^{***} (-96.30)	-0.374 ^{***} (-68.09)	-0.432 ^{***} (-77.16)
manufacturing	-1.183 ^{***} (-299.47)	-1.091 ^{***} (-202.86)	-1.220 ^{***} (-226.87)
maker-dealer	0.973 ^{***} (276.65)	0.388 ^{***} (73.81)	1.314 ^{***} (307.97)
retail	1.104 ^{***} (215.66)	0.598 ^{***} (76.33)	1.436 ^{***} (243.05)
transport	-1.441 ^{***} (-244.10)	-1.963 ^{***} (-197.14)	-1.108 ^{***} (-154.77)
prof and bus services	-0.793 ^{***} (-124.76)	-1.346 ^{***} (-128.37)	-0.432 ^{***} (-56.47)
personal services	0.285 ^{***} (57.55)	-0.697 ^{***} (-67.27)	0.693 ^{***} (123.28)
agric produce processing and dealing	0.113 ^{***} (12.74)	0.0391 ^{***} (3.35)	0.154 ^{***} (13.11)
food sales	1.671 ^{***} (422.80)	0.890 ^{***} (149.74)	2.079 ^{***} (446.76)
refreshment	1.097 ^{***} (217.76)	-0.321 ^{***} (-30.72)	1.646 ^{***} (298.99)
finance and commerce	-0.0181 [*] (-2.10)	-0.561 ^{***} (-41.92)	0.333 ^{***} (33.09)
Constant	-5.454 ^{***} (-585.47)	-6.835 ^{***} (-396.54)	-5.923 ^{***} (-563.73)
Observations	10,637,079		10,637,079
Pseudo R^2	0.279		0.264

Table 3. 1911: Weighted logit model for the probability of being an entrepreneur (employer + own-account) (col.2), and weighted multinomial logit model for the probability of being a worker, or either employer or an own-account self-employed (cols. 3 and 4). Other definitions as in Table 1.

	All entrepreneurs	Employer	Own account
Density	-0.00990*** (-154.83)	-0.00944*** (-98.42)	-0.0102*** (-136.07)
Density # Density	0.0000387*** (100.41)	0.0000376*** (64.85)	0.0000393*** (88.40)
Age	0.158*** (398.38)	0.169*** (258.57)	0.153*** (331.43)
Age # Age	-0.00124*** (-288.08)	-0.00138*** (-199.29)	-0.00118*** (-232.50)
Single men	-0.488*** (-147.45)	-0.740*** (-146.49)	-0.316*** (-78.28)
Widowers	-0.361*** (-73.94)	-0.510*** (-77.14)	-0.234*** (-38.70)
Single women	-0.280*** (-81.11)	-1.300*** (-170.54)	0.0901*** (23.32)
Married women	0.0401*** (9.34)	-0.815*** (-94.16)	0.390*** (82.45)
Widows	0.272*** (57.64)	-0.0435*** (-6.05)	0.450*** (84.03)
Number of servants	0.875*** (284.56)	1.223*** (315.99)	0.425*** (124.39)
mining and quarrying	-2.725*** (-276.90)	-2.635*** (-209.50)	-2.874*** (-183.82)
construction	-0.355*** (-87.91)	-0.393*** (-75.55)	-0.321*** (-57.67)
manufacturing	-1.225*** (-321.80)	-1.135*** (-229.53)	-1.293*** (-237.66)
maker-dealer	0.858*** (248.04)	0.381*** (78.75)	1.204*** (276.16)
retail	0.943*** (210.06)	0.475*** (71.65)	1.293*** (237.30)
transport	-1.660*** (-276.85)	-1.980*** (-220.05)	-1.388*** (-178.34)
prof and bus services	-0.890*** (-164.25)	-1.192*** (-150.87)	-0.624*** (-87.98)
personal services	0.113*** (23.57)	-0.726*** (-82.39)	0.551*** (97.51)
agric produce processing and dealing	0.133*** (15.25)	0.0584*** (5.26)	0.198*** (16.61)
food sales	1.513*** (394.63)	0.949*** (179.21)	1.908*** (404.22)
refreshment	0.826*** (169.18)	0.291*** (38.47)	1.222*** (211.86)
finance and commerce	-0.596*** (-77.03)	-0.907*** (-82.60)	-0.333*** (-33.32)
Constant	-5.556*** (-594.76)	-6.329*** (-411.56)	-6.272*** (-568.52)
Observations	12,071,518		12,071,518
Pseudo R^2	0.263		0.240

Turning to the sector variables in the model we would expect important differences due to market conditions for waged labour that resulted in strong contrasts between mining, heavy manufacturing and transport (dominated by large railway companies and dock enterprises) that employed large numbers of waged personnel on the one hand, intermediate sectors such as construction, and finance and commerce that had a mix of large businesses and smaller traders, and on the other hand sectors such as maker-dealers, refreshments and retail that were predominantly small traders. Hence the main contrasts will reflect sectoral firm-size distribution in the economy as a whole. This is indeed the case. For employers (column 3 of the Tables) the lowest probability of being an employer is in mining, followed in rank order by transport, professional and business services (because of the high level of clerical activity), and then manufacturing. In contrast the highest probability of being an employer is in food sales, followed by more general retail, and then maker-dealers. A similar general grouping of sector probabilities also characterises own account (column 4 of the Tables). But there are some important contrasts. The highest entrepreneurial probability of all is in food sales, followed by refreshment, retail, then maker dealers and personal services. These all reflect ease of market entry and the influence of gender: they are all sectors that were accessible and open to women, offering important opportunities to develop entrepreneurship, for women especially where they otherwise were blocked out of many waged labour markets other than domestic service. The overall pattern of entrepreneurship (column 2) also reflects the own account opportunities and female participation. Food sales, retail, refreshment and maker-dealers offered the greatest opportunities for small business development, especially for women, whilst mining, transport and manufacturing were dominated by waged labour and large firms with few opportunities to develop as employers or own account. The sector estimates therefore demonstrate how different market conditions with varied levels of business concentration interacted with gender as key drivers of choice between employer, own account and worker status.

For the remaining three variables in the estimates (number of servants, density and age) we find a number of important further interactions with gender and sector. The generally strong positive and highly significant interrelationship of entrepreneurship with the number of family servants is to be expected, as is the much higher impact for employers than own account. The number of servants is a composite variable indicating something of the wealth and also other resources available that can release time and may contribute to an entrepreneur's business by servants assisting domestically and in the business, especially in

shops and on farms, although named as servants in the census. Hence employers in many cases, when recording their employer status, may be referring to their employment of staff that were assisting in the business. Because of the joint effects we would expect employers to have a stronger positive relationship with number of servants than own account who are not employers (other than of true domestic servants). Indeed own account and waged workers may be similar. This is borne out in the estimates: employers who are the most strongly differentiated by servant numbers from workers. However, own account, although having a much lower coefficient than for employers, are still strongly differentiated from workers by the servant relationship. Recalling that the census was only assessing full time status of people's occupations, it is likely for own account there is still a joint use of servants as both domestic and business staff, but this reflects a much smaller servant contribution perhaps at a more minor part-time level as far as business contribution was concerned. Given the constraints of the data, however, this is purely speculative; what we can interpret confidently is that number of servants strongly differentiates employers from workers, and at a lesser level than for own account.

Age is estimated as a nonlinear effect as both a direct and squared term. The estimates show that this fits the model found in most literature: that the probability of being an entrepreneur increases with age until a certain point where it levels off and then declines. This differs significantly from worker status. The propensity to be an employer rises more steeply with age than for own account, but the differences are small; the nonlinear effect of declining propensity at higher ages is similar between both types of entrepreneur. The graph for the nonlinear propensity for all entrepreneurs is shown in Figure 1. This also shows the differences between different types of location in terms of level of urbanisation (defined in WP 6). These effects also suggest the influence of additional family farm labour and increased household income needs until dependent children could share decision-making as partners, or left home to develop on their own (Anderson 1971; Davidoff, 2012). The decrease at older ages suggests some effect of succession as older heads withdraw in favour of other family members. The effect of age is in line with modern studies that suggest entrepreneurship is most strongly developed in middle years but the interaction with family needs indicates both resource and sociological interpretations (in-house family available and need). This indicates an important mix of drivers between necessity 'push' and entrepreneurship discovery and opportunism.

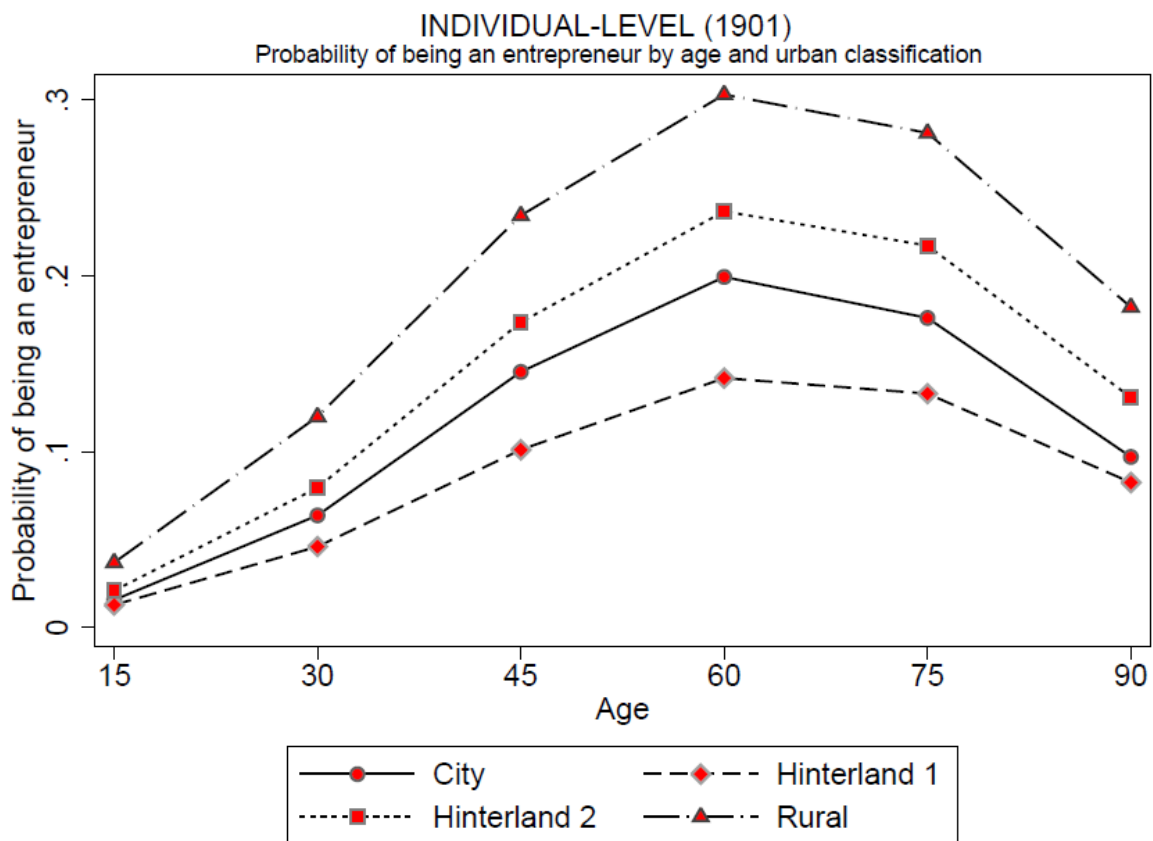


Figure 1. Probability of being either an employer or own account with population density, estimated separately for rural, urban and hinterland areas (hinterland 1 contains non-urban RSDs that contain urban parishes not allocated to a town; hinterland 2 are RSDs that had a population density of more than 0.3 persons per acre; rural are those RSDs with density of less than 0.3 per acre: see WP 6).

Population density is also estimated as a nonlinear effect. The relationship with entrepreneurship is complex. There is generally a declining propensity to be either an employer or own account as density increases, but this is attenuated by the effect of the positive coefficient on the squared term which leads to a higher propensity at higher densities. This indicates behaviourally that much entrepreneurship is highly localised and serves immediate market opportunities, especially at very low population densities in rural areas. But as density increases the opportunities of wider markets beyond the locality, especially in higher urban density areas increases the propensity to be an entrepreneur. As with age, the differences between employers and own account are small. The own account

propensity is slightly higher at both lower and higher densities reflecting their generally greater numbers, but also reflecting a small advantage for easier market entry in both rural and urban conditions. Linking this effect back to the gender structures, this can be interpreted as mainly reflecting the pursuit of own account trading by women as adjoints to their household roles, which tends to restrict trading opportunities to localised markets in both rural and more urban areas.

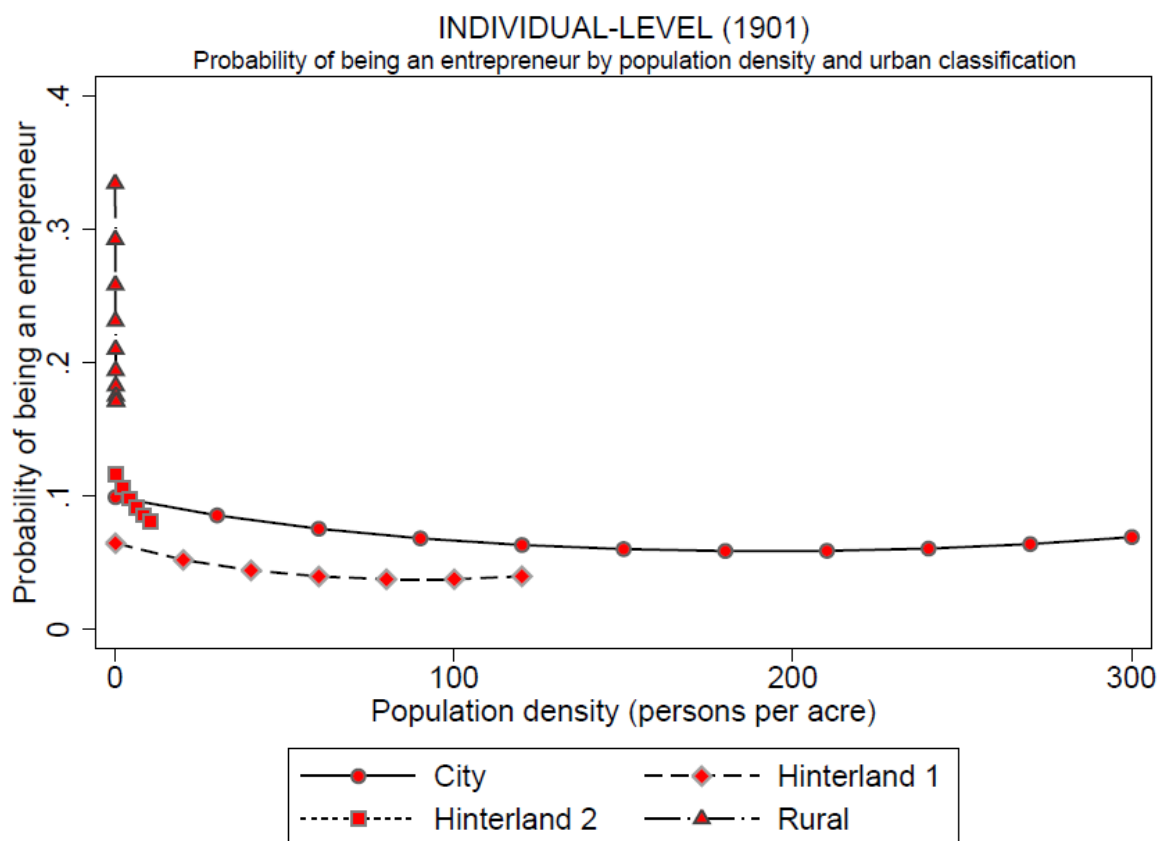


Figure 2. Probability of being either an employer or own account with population density, estimated separately for rural, urban and hinterland areas.

The contrasts between rural and urban areas are drawn out more clearly in Figure 2. This reports the probability levels at different densities after estimating the model separately for different types of location. It shows the generality of the nonlinear effects, but also how these contrast between areas classified as urban, rural, or as different levels of hinterland for large and small urban units. Rural areas have the highest probability of entrepreneurship, but in a very tight locale. This reflects the dominance of the 19th century rural economy by farmers who are all employers or own account, and related businesses. The urban areas have a lower

probability of being an entrepreneur, which reflects the competing opportunities from being waged. But urban entrepreneurship does not vary greatly with density; all urban areas share a level of entrepreneurship that is similar. This in turn is higher than in many hinterland areas of urban centres, where opportunities for being waged are strong and opportunities for entrepreneurship lower than either urban centres or rural areas.

4.2 Model 2: Estimates for interactions of marital status, gender and household structure

Model 1 estimates the effects of marital status and gender as a composite set of variables. Model 2 presents a different approach where the composite categories of gender and marital status are separated between different household relationships using a fully interacted model between gender and marital status (i.e. five levels and six single interactions) for various family and non-family connections identifiable from the census. The estimates are shown in Tables 4-6 for 1891, 1901 and 1911. Worker status is again the base category and in each table there is a comparison between column 2 for the combined probability of being either an employer or own account, and columns 3 and 4 for the probability separately of being an employer or own account compared to being a worker. Apart from the family relationship variables, the other variables (age, density, number of family servants, sector) are the same as in Tables 1-3. For these variables, the estimates are very similar between Tables 1-3 and Tables 4-6 and the same interpretations can be drawn. Hence, we focus attention here on the main issue of interest: gender and household relationships. For these variables the base is male, married, and head of household. Each of the dummy variables is shown as an interaction with these categories. Some of the coefficients in this case are additive and operate as a simple algebra. See Table A2 of the appendix for the transformation of six single interactions of gender and marital status to the eleven fully interacted coefficients for 1911 (notice both regressions have family relationship variables in order to be comparable between them).

Table 4. 1891: Weighted logit model for the probability of being an entrepreneur (employer + own-account) (col.2), and weighted multinomial logit model for the probability of being a worker (omitted) or either employer or an own-account self-employed (cols. 3 and 4). Density at RSD level. (Base categories Male (Sex), Married (Marital Status), Head (Relationship to Head) and Farming/Estate Work (13 occupation categories)).

	All entrepreneurs	Employer	Own account
Density	-0.00694 ^{***} (-133.41)	-0.00701 ^{***} (-86.72)	-0.00693 ^{***} (-118.67)
Density # Density	0.0000192 ^{***} (83.78)	0.0000211 ^{***} (61.80)	0.0000184 ^{***} (71.34)
Age	0.131 ^{***} (294.45)	0.138 ^{***} (184.01)	0.129 ^{***} (253.09)
Age # Age	-0.00101 ^{***} (-212.60)	-0.00107 ^{***} (-138.33)	-0.000990 ^{***} (-180.93)
Female	0.600 ^{***} (98.95)	-0.178 ^{***} (-14.99)	0.869 ^{***} (131.28)
Single	0.0676 ^{***} (13.34)	-0.195 ^{***} (-23.26)	0.261 ^{***} (44.60)
Widow/ed	0.0296 ^{***} (5.62)	-0.288 ^{***} (-38.59)	0.258 ^{***} (41.36)
Female # Single	-0.0437 ^{***} (-5.99)	-0.138 ^{***} (-8.70)	-0.189 ^{***} (-23.78)
Female # Widowed	-0.252 ^{***} (-29.33)	0.239 ^{***} (15.57)	-0.507 ^{***} (-52.91)
CFU member	-0.856 ^{***} (-172.05)	-1.132 ^{***} (-117.24)	-0.779 ^{***} (-143.20)
Older generation	-1.137 ^{***} (-84.69)	-1.251 ^{***} (-50.20)	-1.099 ^{***} (-73.53)
Siblings	-0.832 ^{***} (-96.28)	-0.952 ^{***} (-56.15)	-0.762 ^{***} (-82.93)
Other family	-1.200 ^{***} (-89.73)	-1.725 ^{***} (-50.17)	-1.034 ^{***} (-77.52)
Servants	-4.399 ^{***} (-107.31)	-4.318 ^{***} (-44.19)	-4.502 ^{***} (-99.66)
Working title	-3.264 ^{***} (-96.69)	-2.908 ^{***} (-37.36)	-3.386 ^{***} (-90.89)
Lodgers/boarders	-1.182 ^{***} (-190.72)	-1.455 ^{***} (-116.09)	-1.135 ^{***} (-164.19)
Non-household	-1.900 ^{***} (-87.63)	-2.158 ^{***} (-49.14)	-1.834 ^{***} (-74.94)
Unknown	-0.655 ^{***} (-61.60)	-0.422 ^{***} (-22.26)	-0.753 ^{***} (-63.90)
Number of Servants	0.797 ^{***} (300.84)	1.197 ^{***} (341.90)	0.368 ^{***} (122.26)
mining and quarrying	-2.457 ^{***} (-211.92)	-2.308 ^{***} (-153.47)	-2.651 ^{***} (-143.70)
construction	-0.0233 ^{***} (-5.26)	0.0174 ^{**} (2.99)	-0.0491 ^{***} (-8.12)

manufacturing	-0.879*** (-216.19)	-0.737*** (-134.42)	-0.940*** (-166.20)
maker-dealer	1.145*** (312.02)	0.560*** (103.01)	1.510*** (334.72)
retail	1.387*** (243.15)	0.902*** (107.79)	1.746*** (265.46)
transport	-1.099*** (-175.16)	-1.609*** (-153.80)	-0.731*** (-95.70)
prof and bus services	-0.563*** (-82.19)	-1.074*** (-104.12)	-0.169*** (-19.76)
personal services	0.479*** (92.29)	-0.422*** (-42.72)	0.909*** (150.67)
agric produce processing and dealing	0.382*** (42.59)	0.453*** (40.26)	0.288*** (22.49)
food sales	1.888*** (440.62)	1.157*** (184.26)	2.321*** (458.43)
refreshment	1.342*** (231.44)	0.206*** (20.11)	1.923*** (300.10)
finance and commerce	0.00277 (0.31)	-0.459*** (-35.71)	0.360*** (33.08)
Constant	-4.960*** (-468.17)	-5.750*** (-321.27)	-5.669*** (-463.39)
Observations	8929277	8929277	8929277
Pseudo R ²	0.284		0.268

Table 5. 1901: Weighted logit model for the probability of being an entrepreneur (employer + own-account) (col.2), and weighted multinomial logit model for the probability of being a worker (omitted), or either employer or an own-account self-employed (cols. 3 and 4). Other definitions as in Table 4.

	All entrepreneurs	Employer	Own account
Density	-0.00932*** (-169.30)	-0.00817*** (-89.99)	-0.00975*** (-158.80)
Density # Density	0.0000275*** (100.82)	0.0000245*** (53.65)	0.0000286*** (95.11)
Age	0.144*** (341.31)	0.158*** (206.31)	0.138*** (295.51)
Age # Age	-0.00112*** (-247.33)	-0.00131*** (-163.90)	-0.00104*** (-206.08)
Female	0.715*** (125.21)	0.0424*** (3.57)	0.870*** (142.08)
Single	0.0929*** (19.60)	0.178*** (23.02)	0.151*** (27.94)
Widow/ed	-0.0323*** (-6.28)	-0.134*** (-18.55)	0.0540*** (8.93)
Female # Single	-0.213*** (-31.51)	-0.699*** (-44.54)	-0.223*** (-30.54)

Female # Widowed	-0.317 ^{***} (-38.44)	-0.00207 (-0.14)	-0.403 ^{***} (-44.21)
CFU member	-0.694 ^{***} (-149.63)	-1.464 ^{***} (-152.73)	-0.493 ^{***} (-98.21)
Older generation	-1.265 ^{***} (-97.65)	-2.066 ^{***} (-60.61)	-1.084 ^{***} (-78.51)
Siblings	-0.744 ^{***} (-99.47)	-1.249 ^{***} (-76.85)	-0.572 ^{***} (-72.50)
Other family	-1.069 ^{***} (-84.19)	-2.378 ^{***} (-54.37)	-0.767 ^{***} (-62.02)
Servants	-4.226 ^{***} (-118.02)	-3.863 ^{***} (-58.06)	-4.374 ^{***} (-103.27)
Working title	-2.676 ^{***} (-93.05)	-1.378 ^{***} (-31.77)	-3.106 ^{***} (-83.41)
Lodgers/boarders	-1.256 ^{***} (-205.67)	-1.644 ^{***} (-135.27)	-1.154 ^{***} (-170.16)
Non-household	-1.625 ^{***} (-106.76)	-2.321 ^{***} (-59.59)	-1.481 ^{***} (-89.97)
Unknown	-0.597 ^{***} (-60.52)	-0.179 ^{***} (-11.57)	-0.727 ^{***} (-65.00)
Number of Servants	0.877 ^{***} (303.47)	1.328 ^{***} (345.59)	0.479 ^{***} (149.44)
mining and quarrying	-2.701 ^{***} (-256.25)	-2.587 ^{***} (-185.13)	-2.858 ^{***} (-177.47)
construction	-0.406 ^{***} (-97.92)	-0.371 ^{***} (-67.39)	-0.445 ^{***} (-79.36)
manufacturing	-1.191 ^{***} (-300.87)	-1.084 ^{***} (-201.77)	-1.238 ^{***} (-229.65)
maker-dealer	0.965 ^{***} (272.90)	0.387 ^{***} (73.57)	1.297 ^{***} (302.19)
retail	1.085 ^{***} (210.79)	0.580 ^{***} (73.75)	1.412 ^{***} (237.83)
transport	-1.424 ^{***} (-240.70)	-1.945 ^{***} (-195.30)	-1.094 ^{***} (-152.28)
prof and bus services	-0.796 ^{***} (-125.78)	-1.334 ^{***} (-129.11)	-0.443 ^{***} (-57.74)
personal services	0.372 ^{***} (73.30)	-0.679 ^{***} (-64.76)	0.797 ^{***} (139.31)
agric produce processing and dealing	0.0940 ^{***} (10.65)	0.0191 (1.64)	0.132 ^{***} (11.28)
food sales	1.679 ^{***} (419.88)	0.887 ^{***} (148.42)	2.087 ^{***} (443.83)
refreshment	1.150 ^{***} (219.93)	-0.320 ^{***} (-30.16)	1.728 ^{***} (303.63)
finance and commerce	-0.0243 ^{**} (-2.81)	-0.568 ^{***} (-42.84)	0.323 ^{***} (31.97)
Constant	-5.098 ^{***} (-506.95)	-6.017 ^{***} (-330.39)	-5.721 ^{***} (-502.89)
Observations	10637079		10637079
Pseudo R^2	0.294		0.278

Table 6. 1911: Weighted logit model for the probability of being an entrepreneur (employer + own-account) (col.2), and weighted multinomial logit model for the probability of being a worker (omitted,) or either employer or an own-account self-employed (cols. 3 and 4). Other definitions as in Table 4.

	All entrepreneurs	Employer	Own account
Density	-0.00984*** (-151.99)	-0.00937*** (-96.90)	-0.0101*** (-133.94)
Density # Density	0.0000378*** (97.04)	0.0000365*** (62.23)	0.0000386*** (85.98)
Age	0.144*** (341.06)	0.141*** (206.21)	0.146*** (296.26)
Age # Age	-0.00112*** (-247.49)	-0.00112*** (-157.41)	-0.00112*** (-210.12)
Female	0.605*** (109.67)	0.0962*** (9.82)	0.802*** (130.31)
Single	0.173*** (39.74)	0.150*** (23.04)	0.256*** (49.43)
Widow/ed	-0.0757*** (-14.89)	-0.197*** (-28.82)	0.0393*** (6.31)
Female # Single	-0.441*** (-67.34)	-0.706*** (-56.51)	-0.450*** (-61.51)
Female # Widowed	-0.0899*** (-10.82)	0.256*** (19.30)	-0.249*** (-26.18)
CFU member	-0.729*** (-163.21)	-1.241*** (-158.34)	-0.521*** (-102.96)
Older generation	-1.249*** (-90.41)	-1.850*** (-63.50)	-1.041*** (-68.43)
Siblings	-0.722*** (-101.11)	-1.090*** (-81.01)	-0.549*** (-70.26)
Other family	-0.961*** (-83.22)	-1.641*** (-62.28)	-0.684*** (-56.44)
Servants	-4.126*** (-113.36)	-3.894*** (-63.60)	-4.259*** (-94.45)
Working title	-2.617*** (-106.75)	-1.675*** (-48.55)	-3.114*** (-90.80)
Lodgers/boarders	-1.104*** (-191.12)	-1.313*** (-134.57)	-1.023*** (-152.56)
Non-household	-1.693*** (-116.70)	-2.596*** (-75.06)	-1.435*** (-90.51)
Unknown	-0.581*** (-64.97)	-0.249*** (-19.50)	-0.743*** (-69.13)
Number of servants	0.817*** (270.72)	1.179*** (304.98)	0.351*** (101.76)
mining and quarrying	-2.747*** (-279.20)	-2.657*** (-211.24)	-2.899*** (-185.36)
construction	-0.367*** (-90.79)	-0.401*** (-76.77)	-0.337*** (-60.58)

manufacturing	-1.236*** (-324.62)	-1.135*** (-229.96)	-1.314*** (-241.25)
maker-dealer	0.838*** (241.50)	0.365*** (75.40)	1.175*** (269.01)
retail	0.914*** (202.85)	0.445*** (66.95)	1.259*** (230.50)
transport	-1.652*** (-275.17)	-1.973*** (-219.19)	-1.381*** (-177.23)
prof and bus services	-0.905*** (-167.47)	-1.202*** (-153.52)	-0.646*** (-90.78)
personal services	0.173*** (35.17)	-0.721*** (-80.85)	0.637*** (110.97)
agric produce processing and dealing	0.106*** (12.14)	0.0308** (2.78)	0.168*** (14.15)
food sales	1.500*** (388.48)	0.929*** (174.95)	1.895*** (398.98)
refreshment	0.888*** (176.20)	0.305*** (39.85)	1.315*** (222.48)
finance and commerce	-0.612*** (-79.23)	-0.924*** (-84.78)	-0.352*** (-35.11)
Constant	-5.146*** (-508.58)	-5.565*** (-341.88)	-6.040*** (-503.00)
Observations	12071518	12071518	
Pseudo R^2	0.276	0.253	

The separation of the different effects draws out even more clearly the greater probabilities in general of women being entrepreneurs (once servant status is removed), and this is similar to the pattern in Model 1: that own account offers much more accessible opportunities for women than men because of the limited access of women to the waged labour (other than as servants). But this is now tempered by being able to see separately the interaction effects with marriage. Being single in general increases the probability of entrepreneurship, but for females greatly reduces entrepreneurship probability, especially for being an employer. Widow or widowerhood in general reduces the probability of being an employer, but increases the probability of own account. But decease of spouse for females (widows) reduces the probability of entrepreneurship to a greater extent than for men, especially for own account. This extends the interpretation of Model 1 by indicating that own account is particularly strongly developed by married women, which is in turn a reflection of their labour market opportunities within the marital unit where husbands are predominately waged workers. This in turn indicates a strong necessity motivation; that wives were seeking to supplement household income.

Turning to the probabilities of entrepreneurship within the rest of the household, the head is the base category. Compared with heads (of either gender) all other within-household relations have much lower probabilities of being entrepreneurs. As expected this is lowest of all for servants and those with a working title (who are usually ‘assistants’ in the business), followed by boarders and lodgers, and then those with various levels of family and non-family interrelationship. For employers the highest probability of entrepreneurship is among siblings of the head: brothers and sisters sometimes operate in partnership or operate as heads of different businesses and co-habit the same household. Indeed analysis of partnerships as revealed in the 19th census shows that siblings constituted 28% of all recognisable partnerships in non-farm businesses, and 42% on farms, within which brother-brother partnerships were by far the most frequent: 66% and 94% of sibling partnerships, respectively. Sister-sister were next most common off farms, and brother-sister next most common on farms (Bennett, 2016, Tables 6 and 7). However, the sibling probability of entrepreneurship is still much lower than for heads, and all other household relations have even lower probabilities. This indicates that employers are the dominant leading figures in their households: they are head and all others are part of the domestic support to them releasing resource for the employer to focus on business. They may be supported by brothers in some cases, and in a few cases by sisters or other relatives. This offers a rather more nuanced insight into family businesses than much literature. Because this analysis provides whole-population coverage, we can see any household configuration in comparison to all households. Much modern and historical research has had to focus on case studies or samples where the subject of information is the family firm itself not the rest of firm owners or own account. Looking at the employer in the household in comparison to all other households and other employment statuses shows that in general entrepreneurs tend to suppress the involvement of other family in-house members directly as independent entrepreneurs, even as partners: the head is the entrepreneur and others are part of the wider support network that supports them. Of course those not in the same household are not part of this analysis. If they are partners in the same business with the head, or strike off on their own in another business this is not observable within the data. In addition it has to be noted that the results are only cross sections at one point of time. It may be that, after a period, in-house family and others strike out on their own, or progress to take on the business through various succession strategies. But the results indicate that as far as we can evaluate the probability of entrepreneurship is lower for others within the households of employers as long as the head remains in control of the business.

Turning to own account, although again there are no types of individuals within a household that have a higher probability of entrepreneurship than the own account head, and the range of negative probabilities is similarly wide to that of employer-headed households, some within the household have less negative probabilities than for employers. Among these, other members of the nuclear family (continuous family unit) such as children and grandchildren are the most likely to be also own account, followed by siblings, and other family. These relatives all have higher probability of being an entrepreneur than similar relatives in an employer-headed household. However, for all others in the household the probabilities of entrepreneurship are as low as or lower than for employer households.

All these results confirm the role of family resources, especially from wives, but also sons and daughters, to facilitate increased household income either by supporting the head through domestic duties, pursuing external waged employment, and especially for wives by developing own account activity (for those not in domestic service, public administration etc.). Most restricted among women were those who were single women or widows, though among employers, widows were very similar in probability of being an entrepreneur to married men, perhaps often as result of carrying on the business in which they were already *de facto* partners.

4.3 Comparisons across years

The estimated patterns of relationships are remarkably consistent across the years 1891-1911. However, there are some important indications of trends. To properly investigate these requires a full time series analysis which is beyond the space available here. But certain key features stand out.

First, comparing the coefficients over time for Model 1, there is a steady downward trend in the size of coefficients from 1891 to 1911 suggesting that differences in marital status and gender were diminishing compared to married men, though differences remained highly statistically significant and in essentially the same rank order. The marital unit continued to offer the dominant advantage for employer probability, especially for men, but for own account there remained a higher probability for women. Among women the difference between own account single and married diminished to a very small factor.

Second, between sectors the trends in probabilities varied. However, the changes were consistent for each sector over time. Compared to the base case of farming, there was steadily increasing probability of employer status in mining, processing agricultural produce, food sales, and refreshment. In contrast there was steadily decreasing probability of employer status in maker-dealing, retail, and professional services, and an even more negative probability in manufacturing, construction, transport, personal services, finance and commerce. These trends reflect patterns already known in the secondary literature: the former sectors were offering employers expanded opportunities; the latter sectors were experiencing consolidation into larger firms with reduced scope to develop as independent employers. The changes reflect the development of branches in many sectors and multi-stores in retailing, as well as the gradual shift towards large firm incorporation as limited companies reflecting increasing capital demands and increasing pressure for Chandlerian organisational change. However, at this stage these changes were gradual, though their unidirectional trend is clear.

Third, for own account most sectors showed declining possibilities compared to farming as the base category as consolidation squeezed out scope for easy market entry for entrepreneurs of both genders. Mining, manufacturing, maker-dealing, retail, transport, professional services, personal services, food sales, refreshment and finance and commerce all had smaller own account probabilities by 1911 than at the start of the period. Only construction, and processing of agricultural produce differ, and these are inconsistent: they both had lower probabilities of own account activity 1891-1901, but had higher probabilities 1901-11. These general sector declines patterns are again in line with previous research that has observed increasing consolidation as scale economies in many sectors shifted the balance towards larger plants and organisations, both corporate and non-incorporated, multiple stores expanded in the retail and refreshment sectors, and technological development allowed some larger scale personal services business to develop e.g. through stream laundries. However, the shifts in construction and processing of agricultural require further analysis and comparison with other years.

Fourth, the role of the other variables tended to shift in a complimentary direction. The number of servants became first more important 1891-1901, and then less important 1901-1911 on employer status as well as own account, as the consolidation at the more micro level provided less scope to use domestic supports as a competitive resource. This suggests 1901 as about the turning point where business concentration was generally squeezing

entrepreneurship opportunities; whilst at the same time increases in waged opportunities decreased the attraction and increased the risks of entrepreneurship. The effect of age was up then down for employers, indicating a small shift after 1901 towards higher probabilities for younger-aged individuals operating businesses with their own staff; but the trend was consistently up for own account indicating an increasing probability to operate as self-employed into middle and later years. The negative influence of density increased for both employers and own account, and the density squared term increased for employers, but decreased then increased for own account. These changes reflect a generally increasing impact from density (for own account only after 1901), probably as a result of continuing transport improvements and the expansion of urban areas to cover a larger proportion of the population that encouraged agglomeration economies. This was consistent with, and underpinned, business concentration more generally.

For Model 2 the same patterns generally obtain. The role of household relationships which is the main focus of Model 2 indicate increasing probability for women to be entrepreneurs, especially as employers; lesser scope for single people; and diminished probabilities for widows and widowers. The dominance of the head is similar as for Model 1. The scope for CFU family reduced slightly for employers, but became more negative for own account. There were also increased negative probabilities for the older generation, siblings and other family. Hence, generally the household unit became more dominated by supporting the activities of household heads.

5. Assessment and conclusion

This paper assesses the influences on decisions choices of entrepreneur status for a large sample of the business population for 1891-1911, using a whole population analysis to understanding entrepreneur choices. The census gives a very controlled sample since, under the census instructions, these are businesses are supposed to be full time employers or own account operations, which is a contrast to many modern entrepreneurship samples where very small hours of business activity are often included. The empirical evidence indicates that the choice between business forms (own account proprietorships and employing others) varies significantly by age, location (as measured by population density), wealth or other personnel resources (through the surrogate variable of number of servants), with significant sector differences, and major influences from gender, marital status and family structures. The paper

provides new understanding of how organisational choice was influenced by these different factors.

The results strongly confirm earlier historical studies that view gender and marital status as key influences on business involvement, with the Victorian married couple in particular offering an efficient means to share resources, with men either being freed for wage employment, employer or own account activity (e.g. Anderson, 1971; Davidoff, 2012). The results also confirm modern studies that argue for the importance of the family unit (Carter, 2004) and family size as key influences on entrepreneurship (Carter and Ram, 2003). Hence the model strongly confirms that the scope to share human and other resources with a spouse (mainly) as well as other family members increases the likelihood of entrepreneurship compared to the unmarried. The model also confirms that within the household employers are the most advantaged by the marital relationship in terms of recognised entrepreneur status. The support of the household to entrepreneurship in the 19th century was predominantly for the benefit of men. Where women were involved in entrepreneurship, and were most numerous, was in own account activity to support the marital household. This was likely usually to be as a survival or necessity to increase family income rather than develop truly entrepreneurial businesses.

The paper confirms the significance of family and personal networks to business development in the Victorian era which has been highlighted by Anderson, Davidoff and others. Relative to the base category of married, all marital status-gender combinations (single men, single women, married women, widows and widowers) are significantly less likely to engage in entrepreneurship. This is in line with wider findings of the role of gender and marriage in the 19th century family economy. But there was an important contrast. Married men emerge as the leading employers. They had the support of their wives in the home and often in their business. Married men of course also had a socially privileged status in Victorian England and by the emphasis in the census on the activities of the head of household which was sometimes to the detriment of adequately recording the occupational activities of wives and children, especially daughters. As a result, the results indicate something of a composite of the real social advantages of the married male, and the advantages of the married couple. But married women were also very entrepreneurial, mainly as own account businesses, chiefly in sectors with easy market entry in dressmaking, laundry, inn keeping and shops. This indicates a strong entrepreneurial spirit of self-employment

among women as well as men. Women with independent spirit and opportunity could develop as own account traders to supplement family incomes. This is likely to have been mainly as a response to necessity. A surprising result is the very high levels of female participation in own account activity (once the alternative of domestic service is excluded), usually as an adjoint to a husband who was a waged employee. This was often a necessity to provide sufficient income for the family's well-being. The paper shows that organisational choice between employer and own account status for entrepreneurs, and choice between entrepreneurship and waged status fitted certain market opportunities by sector and location in the past, and was supported by family and wider networks. The decision model demonstrates that organisational choices were interrelated with influences of family structure, sector, location and gender. The piloting of models for three historical censuses can be extended in the future to other years and data sources.

Acknowledgements:

This research has been supported by ESRC grant ES/M010953/1 **Drivers of Entrepreneurship and small business**. We are also pleased to acknowledge the benefits of developing the earlier versions of this analysis by Dragana Radicic, as used for portfolio business analysis in Radicic et al. (2017). An earlier version of the paper was presented at the ISBE Annual Conference in Belfast in November 2017 and we acknowledge the useful comments made there.

The census database derives from K. Schürer, E. Higgs, A.M. Reid, E.M Garrett, *Integrated Census Microdata, 1851-1911, version V. 2 (I-CeM.2)*, (2016) [data collection]. UK Data Service, SN: 7481, <http://dx.doi.org/10.5255/UKDA-SN-7481-1>, enhanced, see: E. Higgs, C. Jones, K. Schürer and A. Wilkinson, *Integrated Census Microdata (I-CeM) Guide*, 2nd ed. (Colchester: Department of History, University of Essex, 2015).

The GIS boundary files for RSDs were constructed by Joe Day for the ESRC fertility project directed by Alice Reid:

<http://www.geog.cam.ac.uk/research/projects/victorianfertilitydecline/publications.html>

These used as a starting point the GIS parish files of Satchell, A.E.M., Kitson, P.M.K., Newton, G.H., Shaw-Taylor, L., Wrigley E.A. (2006) *1851 England and Wales census parishes, townships and places*, 2006, ESRC RES-000-23-1579, supported by

Leverhulme Trust and the British Academy; Satchell, A.E.M. (2015) *England and Wales census parishes, townships and places*; which is an enhanced and corrected version of Burton, N, Westwood J., and Carter P. (2014) *GIS of the ancient parishes of England and Wales, 1500-1850*, UKDA, SN 4828; which is a GIS version of Kain, R.J.P., and Oliver, R.R. (2001) *Historic parishes of England and Wales: An electronic map of boundaries before 1850 with a gazetteer and metadata*, UKDA, SN 4348.

References

- Anderson, M. (1971) *Family Structure in Nineteenth Century Lancashire*, Cambridge University Press.
- Anderson, M. (1988) Households, Families and Individuals: Some Preliminary Results from the National Sample from the 1851 Census of Great Britain, *Continuity and Change*, 3(3), 421-38.
- Alsos, G. A., Carter, S. and Ljunggren, E. (2014) Kinship and business: How Entrepreneurial Households Facilitate Business Growth. *Entrepreneurship and Regional Development* 26 (1-2), 97-122.
- Ang, J. S., Cole, R. E. and Lin, J. W. (2000) Agency costs and ownership structure, *Journal of Finance*, 55, 5, 81-106.
- Armstrong, W. A., (1972) 'The Use of Information about Occupations', in E. A. Wrigley, ed., *Nineteenth-Century Society: Essays in the Use of Quantitative Methods for the Study of Social Data*, Cambridge University Press.
- Bennett, R. J. (2016) Interpreting business partnerships in late Victorian Britain, *Economic History Review*, 69(4), 1199-1227.
- Bennett, R.J., Smith, H. and Monteburno, P. (2019) The Population of Non-corporate Business Proprietors in England and Wales 1891–1911, *Business History*, doi.org/10.1080/00076791.2018.1534959
- Booth, C. (1886) Occupations of the People of the United Kingdom 1801–1881, *Journal of the Royal Statistical Society*, 49, 2, 314-435.
- Cameron, A. C. and Trivedi, P. K. (2005), *Microeconometrics: Methods and Applications*, Cambridge University Press, New York.

- Carter, S. and Ram, M. (2003) Reassessing Portfolio Entrepreneurship, *Small Business Economics*, 21(4), 371-80.
- Chandler, A. D. (1962) *Strategy and Structure: Chapters in the History of the Industrial Enterprise*, MIT Press, Cambridge Mass.
- Church, R. (1993) 'The Family Firm in Industrial Capitalism: International Perspectives on Hypotheses and History', *Business History*, 35, 4, 17-43.
- Colli, A. (2003) *The History of Family Business, 1850-2000*, Cambridge University Press.
- Colli, A., Howorth, C. and Rose, M. (2013) Editorial: Long Term Perspectives on Family Business, *Business History*, 55, 6, 841-54.
- Davidoff, L. (2012) *Thicker Than Water: Siblings and their Relationships 1780-1920*, Oxford University Press.
- Davidoff, L. and Hall, C. (1997) *Family Fortunes: Men and Women of the English Middle Class 1780-1850*, Revised Edition, Routledge, London.
- Guo, G., and Zhao, H., (2000) Multilevel Modelling for Binary Data. *Annual Review of Sociology*, 26, 441–462.
- Higgs, E., Jones, C., Schürer, K. and Wilkinson, A. (2015), *Integrated Census Microdata (I-CeM) Guide*, 2nd ed., Colchester: Department of History, University of Essex.
- Howorth, C., Rose, M., Hamilton, E. and Westhead, P. (2010) Family firm diversity and development: An introduction, *International Small Business Journal*, 28, 5, 437-451.
- Jensen, M. C. and Meckling, W. H. (1976) Theory of the firm: Managerial behaviour, agency costs and ownership structure, *Journal of Financial Economics*, 3, 305-360.
- Knight, F.H. (1921) *Risk, Uncertainty and Profit*, Boston, MA: Houghton Mifflin; reissued with additional Prefatory Essay, London School of Economics, 1933.
- La Porta, R., Lopez-de-Silanes, F., Schleifer, A. and Vishny, R. W. (1998) Law and finance, *Journal of Political Economy*, 106, 6, 1113-1155.
- La Porta, R., Lopez-de-Silanes, F. and Schleifer, A. (1999) Corporate ownership around the world, *Journal of Finance*, 54, 2, 471-517.
- La Porta, R., Lopez-de-Silanes, F. and Schleifer, A. (2008) The economic consequences of legal origins, *Journal of Economic Literature*, 42, 2, 285-332.
- Monteburuno, P. (2018) *Research data supporting Adjustment Weights for the 1891-1911 England and Wales censuses* [Dataset]. <https://doi.org/10.17863/CAM.26376>
- Nenadic, S., (1993) The Small Family Firm in Victorian Britain, *Business History*, 35, 86-114.

- Pollard, S. (1965) *The Genesis of Modern Management: A Study of the Industrial Revolution in Great Britain*, Edward Arnold, London.
- Rabe-Hesketh, S. and Skrondal, A., (2012), *Multilevel and Longitudinal Modeling Using Stata*, 3rd Edition, 3rd ed., StataCorp LP.
- Radacic, D., Bennett, R.J. and Newton, G. (2017) Portfolio entrepreneurship in farming: Empirical evidence from the 1881 census for England and Wales, *Journal of Rural Studies*, 55, 289-302.
- Schürer, K., Higgs, E., Reid, A.M. and Garrett, E.M. (2016), *Integrated Census Microdata, 1851-1911, version V. 2 (I-CeM.2)*, [data collection]. UK Data Service, SN: 7481, <http://dx.doi.org/10.5255/UKDA-SN-7481-1>
- Williamson, O. E. (1975) *Markets and Hierarchies*, Free Press, New York.

Appendix:**Table A1. Variable description and summary statistics (weighted for nonresponse bias)**

Variable name	Variable description	Mean (standard deviation)		
		1891	1901	1911
Dependent variables - Employment status	DV= employer, own account, worker (base category)	See below		
Worker	DV=1 if worker	0.87 (0.34)	0.87 (0.33)	0.89 (0.31)
Employer	DV=1 if employer	0.05 (0.21)	0.04 (0.2)	0.04 (0.21)
Own account	DV=1 if own account	0.09 (0.28)	0.09 (0.28)	0.07 (0.25)
Age	Age (as continuous variable)	34.24(15.39)	33.94(15.02)	34.74(14.91)
Density	Population per acre for RSD	30.58(50.38)	29.88(47.83)	29.16(40.42)
Number of servants	No. of domestic servants in household	0.13 (0.58)	0.11 (0.52)	0.1 (0.54)
Single men	DV=1 if a single man (base category)	0.27 (0.45)	0.29 (0.45)	0.27 (0.45)
Single women	DV=1 if an employer single woman	0.22 (0.41)	0.22 (0.42)	0.23 (0.42)
Married men (base category)	DV=1 if a married man	0.38 (0.49)	0.38 (0.49)	0.38 (0.49)
Married women	DV=1 if a married woman	0.05 (0.22)	0.04 (0.19)	0.05 (0.22)
Widowers	DV=1 if a widower	0.03 (0.18)	0.03 (0.17)	0.03 (0.17)
Widows	DV=1 if a widow	0.04 (0.19)	0.03 (0.18)	0.03 (0.16)
Head	The head of household (base category)	0.44 (0.5)	0.43 (0.5)	0.42 (0.49)
CFU member (within nuclear family)	Spouse of head and children	0.3 (0.46)	0.31 (0.46)	0.33 (0.47)
Older generation	Parent of head, grandparent, uncle/aunt	0.01 (0.07)	0.01 (0.07)	0.01 (0.07)
Siblings	Brothers and sisters of head	0.02 (0.15)	0.03 (0.16)	0.03 (0.16)
Other family	Grandchildren of head, cousin, niece/nephew, extended other and miscellaneous relatives	0.02 (0.13)	0.02 (0.13)	0.02 (0.13)
Servants	Domestic live-in servants	0.09 (0.29)	0.08 (0.28)	0.07 (0.26)
Working title	Live-in workers under head	0.01 (0.08)	0.01 (0.07)	0.01 (0.08)
Lodgers & boarders	Those renting long-term at census	0.08 (0.27)	0.08 (0.27)	0.07 (0.26)
Non-household	Where premises are an institution (school, hospital, large hotel,	0.03 (0.14)	0.02 (0.14)	0.02 (0.13)

	military establishment)			
Unknown relationship	All others and visitors at time of census	0.02 (0.14)	0.02 (0.14)	0.03 (0.15)
Farming	DV = 1 for farmers (base category)	0.11 (0.31)	0.08 (0.28)	0.08 (0.27)
Mining and quarrying	DV = 1 for mining and quarrying	0.05 (0.22)	0.05 (0.23)	0.06 (0.24)
Construction	DV = 1 for construction sector	0.06 (0.25)	0.08 (0.27)	0.06 (0.25)
Manufacturing	DV = 1 for manufacturing sector	0.2 (0.4)	0.2 (0.4)	0.2 (0.4)
Maker-dealer	DV = 1 for manufacturing & retailing	0.12 (0.33)	0.12 (0.32)	0.1 (0.3)
Retail	DV = 1 for retail sector	0.02 (0.14)	0.02 (0.15)	0.03 (0.17)
Transport	DV = 1 for transport sector	0.07 (0.25)	0.08 (0.27)	0.08 (0.26)
Professional and business services	DV = 1 for professional services (e.g. lawyers, architects, scientific pursuits)	0.03 (0.18)	0.04 (0.2)	0.05 (0.22)
Personal services	DV = 1 for personal services (e.g. doctors, dentists, artists, performers, education)	0.06 (0.23)	0.07 (0.23)	0.07 (0.23)
Agricultural produce processing and dealing	DV = 1 for agricultural produce	0.01 (0.09)	0.01 (0.09)	0.01 (0.08)
Food sales	DV = 1 for produce processors	0.05 (0.21)	0.05 (0.21)	0.04 (0.2)
Refreshment	DV = 1 for publicans, coffee houses, lodgings, hotels	0.03 (0.15)	0.02 (0.15)	0.03 (0.16)
Finance and commerce	DV= 1 for merchants, bankers, insurers, brokers, accountants, salesmen, etc.	0.01 (0.11)	0.02 (0.11)	0.02 (0.13)
Public admin., military, clergy	Excluded from models	0.02 (0.16)	0.03 (0.17)	0.03 (0.16)
Domestic service	Excluded from models	0.11 (0.32)	0.11 (0.32)	0.1 (0.31)
Undefined and general labourers	Excluded from models	0.05 (0.22)	0.03 (0.18)	0.04 (0.2)
Persons of property	Excluded from models	0 (0.02)	0 (0.02)	0.01 (0.1)

Table A2. Algebra of coefficients (example for 1911 with family relationship variables)

Regressions with six single interactions		Regression with eleven fully interactions.				
Single men 0.17	=	Male 0	+	Single 0.17	+	Male Single 0
Married men 0	=	Male 0	+	Married 0	+	Male Married 0
Widowers -0.08	=	Male 0	+	Widow/ed -0.08	+	Male Widowers 0
Single women 0.34	=	Female 0.60	+	Single 0.17	+	Female Single -0.44
Married women 0.60	=	Female 0.60	+	Married 0	+	Female Married 0
Widows 0.44	=	Female 0.60	+	Widow/ed -0.08	+	Female Widows -0.09

Other Working Papers:

Working paper series: ESRC project ES/M010953: *'Drivers of Entrepreneurship and Small Business'*, University of Cambridge, Department of Geography and Cambridge Group for the History of Population and Social Structure.

WP 1: Bennett, Robert J., Smith Harry J., van Lieshout, Carry, and Newton, Gill (2017) *Drivers of Entrepreneurship and Small Businesses: Project overview and database design*. <https://doi.org/10.17863/CAM.9508>

WP 2: Bennett, Robert J., Smith Harry J. and van Lieshout, Carry (2017) *Employers and the self-employed in the censuses 1851-1911: The census as a source for identifying entrepreneurs, business numbers and size distribution*. <https://doi.org/10.17863/CAM.9640>

WP 3: van Lieshout, Carry, Bennett, Robert J., Smith, Harry J. and Newton, Gill (2017) *Identifying businesses and entrepreneurs in the Censuses 1851-1881*. <https://doi.org/10.17863/CAM.9639>

- WP 4: Smith, Harry J., Bennett, Robert J., and van Lieshout, Carry (2017) *Extracting entrepreneurs from the Censuses, 1891-1911*. <https://doi.org/10.17863/CAM.9638>
- WP 5: Bennett, Robert J., Smith Harry J., van Lieshout, Carry, and Newton, Gill (2017) *Business sectors, occupations and aggregations of census data 1851-1911*. <https://doi.org/10.17863/CAM.9874>
Data download of classification file: <https://doi.org/10.17863/CAM.9874>
- WP 6: Smith, Harry J. and Bennett, Robert J. (2017) *Urban-Rural Classification using Census data, 1851-1911*. <https://doi.org/10.17863/CAM.15763>
- WP 7: Smith, Harry, Bennett, Robert J., and Radicic, Dragana (2017) *Classification of towns in 1891 using factor analysis*. <https://doi.org/10.17863/CAM.15767>
- WP 8: Bennett, Robert J., Smith, Harry, and Radicic, Dragana (2017) *Classification of occupations for economically active: Factor analysis of Registration Sub-Districts (RSDs) in 1891*. <https://doi.org/10.17863/CAM.15764>
- WP 9: Bennett, Robert, J., Montebruno, Piero, Smith, Harry, and van Lieshout, Carry (2018) *Reconstructing entrepreneurship and business numbers for censuses 1851-81*.
- WP 10: Bennett, Robert, J., Smith, Harry and Radicic, Dragana (2018) *Classification of environments of entrepreneurship: Factor analysis of Registration Sub-Districts (RSDs) in 1891*. <https://doi.org/10.17863/CAM.26386>
- WP 11: Montebruno, Piero (2018) *Adjustment Weights 1891-1911: Weights to adjust entrepreneur numbers for non-response and misallocation bias in Censuses 1891-1911*. <https://doi.org/10.17863/CAM.26378>
Adjustment weights: <https://doi.org/10.17863/CAM.26376>
- WP 12: van Lieshout, Carry, Day, Joseph, Montebruno, Piero and Bennett Robert J. (2018) *Extraction of data on Entrepreneurs from the 1871 Census to supplement I-CeM*. <https://doi.org/10.17863/CAM.27488>
- WP 13: van Lieshout, Carry, Bennett, Robert J., and Smith, Harry (2019) *Extracted data on employers and farmers compared with published tables in the Census General Reports, 1851-1881*.
- WP 14: van Lieshout, Carry, Bennett Robert J. and Montebruno, Piero (2019) *Company Directors: Directory and Census record linkage*.
- WP 15: Bennett, Robert, J., Montebruno, Piero, Smith, Harry and van Lieshout, Carry (2019) *Entrepreneurial discrete choice: Modelling decisions between self-employment, employer and worker status*.

Full list of all current Working Papers available at:

<https://www.campop.geog.cam.ac.uk/research/projects/driversofentrepreneurship/>